

DOCUMENT RESUME

ED 318 630

SE 051 379

AUTHOR Lowry, William H.; Thorkildsen, Ron
TITLE The Effects of a Videodisc-Based Direct Instruction Program in Fractions on Mathematics Achievement and Self-Concept.
PUB DATE Apr 90
NOTE 37p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Research/Technical (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Audiovisual Aids; Educational Technology; *Elementary School Mathematics; *Fractions; Grade 4; Grade 5; Intermediate Grades; *Mathematics Achievement; Mathematics Education; *Self Concept; *Videodisks

ABSTRACT

This study investigated the effect of a direct instruction program in fractions, on achievement and self-concept. A pretest-posttest, quasi-experimental, nonequivalent control group design was used to compare the two groups of upper elementary school students (N=337). Treatment group students received instruction in fractions via a teacher-directed, videodisc-based program. The control group received their normal grade-four or grade-five mathematics program. Results indicated that the treatment group mean on a criterion referenced test was higher than that of the control; differences in achievement test scores among the treatment classes varied directly with the level of program implementation; and the scores of the treatment group classes on the self-concept instrument were higher than those of the control group classes. A list of 31 references is included. (Author/CW)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED318630

THE EFFECTS OF
A VIDEODISC-BASED DIRECT INSTRUCTION PROGRAM IN FRACTIONS
ON MATHEMATICS ACHIEVEMENT AND SELF-CONCEPT

by

William H. Lowry and Ron Thorkildsen

Utah State University

Logan, Utah 84322-6800

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

William H. Lowry

for

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

The Annual Meeting
of the
American Educational Research Association

April 1990

The paper is abstracted from Lowry, W. H. (1989). The effects of a direct instruction program in fractions on academic and mathematics self-concept. (Doctoral dissertation, Utah State University). The study was completed as part of a parent project funded by the United States Department of Education, project number 84.180R2--Technology, Educational Media & Materials Assessing the Social and Cultural Impact of a Group-Based, Videodisc Math Instruction in Mainstreamed Classrooms.

The Self-Description Questionnaire (the measure of self-concept used in this study) is a copyrighted publication of The Psychological Corporation, and is used by permission. No reproduction the SDQ may be made without permission of the Publisher.

SE051379

Abstract

The study investigated the effect of a direct instruction program in fractions, on achievement and self-concept. Self-concept was operationally defined as scores on a slightly modified version of Marsh's Self-Description Questionnaire (Psychological Corporation, 1988); achievement as scores on a criterion-referenced test of fractions skills and concepts. We used a pretest-posttest, quasi-experimental, nonequivalent control group design to compare the self-report self-concept of two groups of upper elementary students ($N = 337$). Treatment group students ($n = 171$) received instruction in fractions via the teacher-directed, videodisc-based, Mastering Fractions program (Systems Impact, 1986). The control group ($n = 166$) received their normal grade-four or grade-five mathematics program, which did not include common fractions.

The treatment group covariance-adjusted mean on a criterion-referenced test was higher than that of the control (5.9 standard deviations). These differences in achievement scores provided support for previous findings regarding the Mastering Fractions program. Differences in achievement test scores among the treatment classes varied directly with the levels of program implementation across classes.

The self-concept data were examined using both the student and the class as the unit of analysis. Using the

student as the unit of analysis, the treatment-group, mathematics self-concept, covariance-adjusted mean was 0.22 standard deviations above that of the control group. A statistically significant but small main effect was also noted across student achievement levels. The posttest difference between low-achiever means (treatment versus control students) is slightly larger than the difference between high-achiever means. No statistically significant interaction was noted between student achievement level at pretest and treatment condition. An analysis of raw gain scores yielded a standardized mean difference effect size between the treatment and control group scores of +.12.

The class was also used as the unit of analysis. In this case the mean difference effect sizes between experimental groups were +0.86 (ANCOVA) and +0.31 (raw gain scores).

Introduction

Several authors have suggested that the use of direct instruction should result in positive changes in self-concept and academic achievement (e.g., Engelmann, 1980). Further, the literatures concerning process-product research (see e.g., reviews cited in Waxman and Walberg, 1982), self-efficacy (e.g., Bandura, 1977), and effective teaching (e.g., Rosenshine, 1979) all lend support to the hypothesis that specific components of the direct instruction can promote achievement and self-concept. The Direct Instruction Model (e.g., Engelmann, 1980) is a specific form of direct instruction. Thus far direct instructional approaches, and in particular the total Direct Instruction Model, have been tested for their effect on self-concept in only a few studies.

In the past researchers have found self-concept to be relatively stable over time and that there is a low positive correlation between academic achievement and self-concept (e.g., Marsh, 1988). Prior to the mid-1970s, self-concept among elementary-school children was thought to be dominated by a general self-concept factor (Marsh, 1988). To the contrary, Shavelson and Marsh, and other researchers, have since found support for a hierarchical/multidimensional model of elementary-age student self-concept (see Shavelson, Hubner, & Stanton, 1976). The SDQ contains separate scales

to measure the various dimensions of self-report self-concept.

The Direct Instruction Model is achievement-oriented rather than self-concept-oriented. Thus, one of the objectives of the study was to measure fractions achievement, to test whether the treatment actually taught fractions, before we drew any conclusions regarding self-concept. To meet this objective, we employed a criterion-referenced, domain-specific test in fractions which is a part of the Mastering Fractions program.

Procedures

The portion of the study reported in this paper was conducted during the fall of 1988 in northern Utah, northern Wyoming, and southern New Mexico. Students who were in the control group during the fall study received the program during winter and spring of 1989.

Subjects

The sample contained primarily grade-five classes, but included two grade-four classes. On achievement, an important variable in this study, the students were comparable to their national grade-level cohorts. The sample data did differ from national census data on ethnic origin (37.7% Hispanic origin) and slightly on age.

It was our intention to randomly assign treatment condition to classes, but we were unable to do so because

some teachers were unwilling to teach fractions during fall quarter. Consequently, these classes became part of the control group. We requested permission for the 434 students to respond to the self-concept measure; and received permission for 399 (91.9%) students to participate. During the course of the study, 62 subjects were lost because of transfer and illness during testing. The resulting sample included 337 students: 171 in the treatment group and 166 in the control group.

Research Design

We used a quasi-experimental, pretest-posttest, non-equivalent control group design to investigate the effect of a videodisc-based, teacher-controlled, Direct Instruction Model-based program in fractions on achievement and self-concept. All classes received the pretests on achievement and self-concept just before the experimental group began the Mastering Fractions program. Pretests were conducted by site coordinators. These test administrators were provided written protocols for test administration. Protocols were particularly important for the SDQ-I, as they provided some protection against fakability. In addition, students were informed that their teachers would not see the completed instrument and were encouraged to cover their responses.

Observations of the implementation of the treatment were made throughout the fall, as the students progressed through the program. When experimental classes at a given

site had completed the Fractions program, all classes at that site were posttested on achievement and self-concept. Four to six weeks after the completion of the program, follow-up measures were administered to all experimental-group students.

Instructional Programs

The Mastering Fractions program is designed to be used with a heterogeneous class of students. The program developers provide recommendations for program use. A large television or monitor (screen-size large enough to be read in all parts of the classroom) is placed at the front of the class. All students are seated facing the screen. A consumer model videodisc player is connected to the television and is controlled by the teacher from anywhere in the classroom through the use of a remote control unit. The program requires the Mastering Fractions videodiscs, consumable student worksheets, extra paper and pencils, and the teacher handbook (which includes answer keys for the seat work and tests).

Mastering Fractions consists of 35 lessons. Each lesson is designed to be completed in approximately 40 minutes. Most lessons begin with a quiz covering the previous lesson's learning, and each fifth lesson is a mastery test. Both quizzes and mastery tests are followed by diagnostic/remedial suggestions based on specific skill weaknesses manifested by the students. Narrations and

visuals recorded on the videodisc program present the instruction; and students interact with the program through oral response and written practice.

Each presentation/practice lesson provides concept and skill teaching, wherein the students' understanding is checked by the quality of their choral responses. If the students fail to provide an adequate response to the program's questions, it is recommended that the teacher encourage response or provide remediation. Each concept or component skill is designed to be a small step toward a problem-solving strategy.

If most of the students are successful in their oral responses, the teacher allows the program to proceed to written, guided-practice. Students are asked to solve problems. The teacher checks to see if 80% of the students were successful on each written component skill, by checking a show of hands or by circulating among the students. If the criterion is not met, the teacher branches to a specified part of the videodisc for a remedial presentation.

If written practice is successful, students move on to independent practice that integrates skills and concepts. It is recommended that independent practice be done as seat work with the teacher circulating, checking progress, and providing guidance to students experiencing difficulties.

Teachers are encouraged to be positive, to maintain a brisk pace, and to verbally reward success. The diagnostics

provided by the program are used to pace the overall program to student learning needs. However, even the pace of remedial work and help at seat work is to be brisk. That is, the pace of the work reflects student learning needs, not student desires.

We provided teachers with training in the use of the program, as recommended by the program developers. Program implementation was measured against guidelines presented during that training.

Teachers in the nonfractions control group used the math materials they normally used. These materials covered whole-number and decimal concepts and skills. Activities included lectures, games, in-class seat work, quizzes, etc. Control group teachers did not systematically teach fractions since they would be using the Mastering Fractions program during winter/spring of the same academic year.

Achievement Test Scores

We used a criterion-referenced (domain-referenced) test, developed for previous research with Mastering Fractions, as a measure of achievement. The test contains 57 items covering fractions skills and concepts. Since it is intended to be criterion based, there were low scores and low variability among scores at pretest. The test-retest correlation for control students is +0.67. Pretest scores from the criterion-referenced test correlate at +0.56 with percentage correct scores on norm-referenced tests (the ITBS

or CTBS), which is understandable considering the floor effect of the criterion test.

Observations

Implementation guidelines for Mastering Fractions are provided in the paper-based support materials supplied with the program. The guidelines, along with suggestions made by observers experienced in research on the program, and ideas derived from watching teachers use the program were combined in a printed list of suggestions for implementation which was given to the teachers and discussed during in-service training. The observation form we used was a copy of the list of suggestions. Exact interobserver agreement was 75.6%; interobserver agreement, defined as "within one step on the rating scale," was 96.7%.

Self-Concept Ratings

We chose the SDQ-I, because it is based on current research findings regarding the self-concept construct, and because it fit the needs of the study. Several researchers have conducted studies with grade-five students to test the reliabilities of the instrument and its scales (Marsh, Smith, Barnes, & Butler, 1983; Marsh, 1988). Cronbach alpha values and test-retest correlations reported in the SDQ Manual (Marsh, 1988, p. 45) are presented in Table 1 and compared to the values from the present study. Academic self-concept is operationally defined as the average of

Table 1

Reliability Correlation Coefficients

	School Self-Concept	Reading Self-Concept	Math Self-Concept
Cronbach Alpha <u>SDQ Manual</u>	0.86	0.89	0.89
Cronbach Alpha present study	0.84	0.89	0.91
Test-retest Correlation (6 months) <u>SDQ Manual</u>	0.54	0.61	0.68
Test-retest Correlation present study (20 days)	0.86	0.82	0.88
Test-retest Correlation present study (90 days)	0.64	0.51	0.62

school, reading, and mathematics self-concepts. Marsh notes that the moderate test-retest correlations over a six-month period indicate that (while self-concept is relatively stable over time) some systematic changes in self-concept do occur during the school year (Marsh, 1988; Marsh, Smith, Barnes, & Butler, 1983).

Tests of construct validity of the SDQ-I have been conducted on a number of samples testing convergent and discriminant validity of the scales, with good results (Marsh, Relich, & Smith, 1983; Marsh & Smith, 1987; Marsh, Smith, Barnes, & Butler, 1983). The SDQ-I has also been

tested against teacher ratings of student self-concept (Marsh, Parker, & Smith, 1983; Marsh, Smith, & Barnes, 1981, 1983, 1984). Most of the development work has been done with Australian youth; however, more recent work has been done in England (Marsh & Smith, 1987). A longitudinal study is underway in the United States that involves the SDQ; results will be available soon (A. Gottfried, personal communication, June 25, 1989).

To determine the acceptability of the instrument for American students, we submitted the complete SDQ-I to a local principal and several elementary school teachers. Each individual was asked to provide impressions about the instrument. Of the seven original scales, the present study was concerned only with mathematics, reading, and school-related self-concepts. Several items within those scales, that contained the word "dumb" (numbers 33, 47, & 72) were seen as potentially offensive to parents or students, although neither the principal nor the teachers recommended that the scale be modified. A fellow staff member with public school experience, who was planning to use the SDQ-I, also objected to these items. The same items were questioned in Marsh's cross-cultural study in England (Marsh & Smith, 1987).

The negatively-cast items on the SDQ-I (e.g., "I am dumb at mathematics"; "I hate mathematics.") are not scored. Rather they are included in the instrument to prevent response set (Marsh, 1988). We contacted The Psychological

Corporation (current owner of the rights to the SDQ) to gain permission to delete or modify the challenged items. We removed only the items which were seen as potentially offensive. The resulting set of school-related scales includes 35 items. Each positively-cast item was rated on a one-to-five scale, with one as low (Marsh, 1988). The rating was then determined for each scale by summing the individual item ratings on the particular scale. The maximum scale score is 40.00 rating points, the minimum is 00.00.

Results and Discussion

The main body of this section of the report is organized by hypotheses. The section immediately following, however, deals with objectives of the study--to measure achievement and to determine level of implementation of the Mastering Fractions program across classes.

Achievement Differences and Program Implementation

Some hypotheses regarding self-concept required the measurement of math achievement. Differences in achievement scores provided support for previous findings regarding the Mastering Fractions program (e.g., Hasselbring et al., 1986). The treatment group covariance adjusted mean on a criterion-referenced test was higher than that of the control (5.9 standard deviations). The difference is statistically significant at less than the .001 level.

The effect of levels of implementation is statistically significant as well ($F = 25.92$; $p < .001$) (see Table 2), which is also consistent with findings made by Hasselbring et al. (1986). Note that there is a substantial difference in achievement scores across levels of implementation.

Table 2

Summary of Adjusted Cell Means for the Analysis of Covariance

Level of Implementation	Mean Achievement Percentage	n
High	83.6%	55
Average	76.1%	54
Low	63.9%	62

Self-Concept Differences

Hypothesis 1. Within the treatment group, the mean mathematics self-concept gain rating, pre- to posttest, is significantly greater than the mean academic self-concept gain rating. (Academic self-concept is defined as the average of the reading, math, and school self-concept scale ratings.)

Table 3 contains the descriptive statistics of gain ratings for the two self-concept scales. Although the difference is in the predicted direction, the gains in both cases are small.

Table 3

Means, Standard Deviations, Correlation Coefficients,
T Test results, and the Effect Size for Math Self-
Concept and Academic Self-Concept Gain Scores
for the Treatment Group

Instrument	$\bar{X}$	Gain SD	n-size
Academic Self-Concept	0.55	4.37	171
Math Self-Concept	0.89	6.14	171
$r = +0.74$			
$t = +1.08 \quad p = 0.14$ (directional test)			
Standardized Mean Difference Effect Size = +0.05 (using the pooled pretest standard deviation of raw ratings)			

Considering math self-concept alone, there is a 0.89 point mean gain out of possible 10 point gain (based on pretest ratings of around 30 points and a ceiling of 40 points). If the math self-concept scale is an equal-interval scale, this would represent about a 9% gain relatively near the ceiling of the instrument. If the SDQ math scale had not had a ceiling of 40 points, the gain might have been more substantial.

Table 4 presents statistical information for the control group data for comparison. In the case of the control group there was essentially no change in math self-concept, while the gain in academic self-concept ratings is comparable to that in the treatment group.

Table 4

Means, Standard Deviations, Correlations Coefficients, T Test results, and the Effect Size for Math Self-Concept and Academic Self-Concept Change Scores for the Control Group

Instrument	$\bar{X}$	SD	n-size
Academic Self-Concept	0.52	5.01	166
Math Self-Concept	-0.04	7.11	166

$r = +0.76$
 $t = -1.57$ $p = 0.06$ (directional test)
 Standardized Mean Difference Effect Size = -0.08
 (using the pooled pretest standard deviation of raw ratings)

Hypothesis 2. The mean posttest rating of the treatment group mathematics self-concept measure is significantly greater than the mean posttest rating of the control group.

Hypothesis 3. There is a significant interaction between mathematics self-concept ratings and levels of student achievement (high, medium, and low--based on pretest scores on a standardized mathematics achievement test).

To test these two hypotheses, we used analysis of covariance procedures, first using "student" as the unit of analysis, then using "class" as the unit of analysis. To test hypothesis 2 we also considered a gain score analysis.

Analyses using "student" as the unit of analysis.

Table 5 presents the unadjusted descriptive statistics for

the comparison. The breakdown is by pretest achievement level.

Table 5

Unadjusted Means, Standard Deviations, and N-sizes
for the Pre- and Posttest Measures of Math Self-Concept--
Student as the Unit of Analysis

Instructional Method by Achievement Level	Pretest		Posttest		n-size
	$\bar{X}$	SD	$\bar{X}$	SD	
<u>Mastering Fractions Group</u>	30.79	7.77	31.68	7.06	171
High	32.65	6.83	33.44	6.45	57
Middle	30.68	7.66	31.56	6.40	57
Low	29.03	8.43	30.05	7.94	57
<u>Control Group</u>	28.88	8.54	28.84	7.72	166
High	31.20	8.27	31.54	7.57	55
Middle	30.30	7.87	29.80	6.94	55
Low	25.11	8.32	25.14	7.31	56

Tables 6 and 7 summarize a 2 X 3 analysis of covariance, with math self-concept pretest scores as the covariate. Both main effects (for achievement group membership and treatment group membership) are statistically significant (achievement group-- $F = 4.09$, $p < .02$; treatment group-- $F = 8.41$, $p < .01$).

Table 6

Analysis of Covariance for Posttest Math Self-Concept
using Pretest Math Self-Concept as the Covariate--Student
as the Unit of Analysis

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F	Significance of the F
Within cell	10526.82	330	31.90		
Regression	6263.42	1	6263.42	196.35	<.001
Constant	4180.90	1	4180.90	131.06	<.001
RGROUP*	268.22	1	268.22	8.41	.004
ACH**	260.83	2	130.41	4.09	.018
RGROUP BY ACH	40.57	2	20.28	.64	.530

* RGROUP = Treatment group membership
 ** ACH = Achievement level

Table 7

Summary of Adjusted Math Self-Concept
Rating Cell Means for the Analysis of
Covariance--Student as the Unit of Analysis

Research Group	Pretest Achievement Level			Main Effect
	High	Middle	Low	
Treatment	31.89	31.09	30.49	31.16
Control	30.79	29.54	27.74	29.36
Main Effect	31.34	30.32	29.12	

A graphical representation of the data broken down by achievement group (figure below) does show some convergence toward the high-achiever end of the graph.

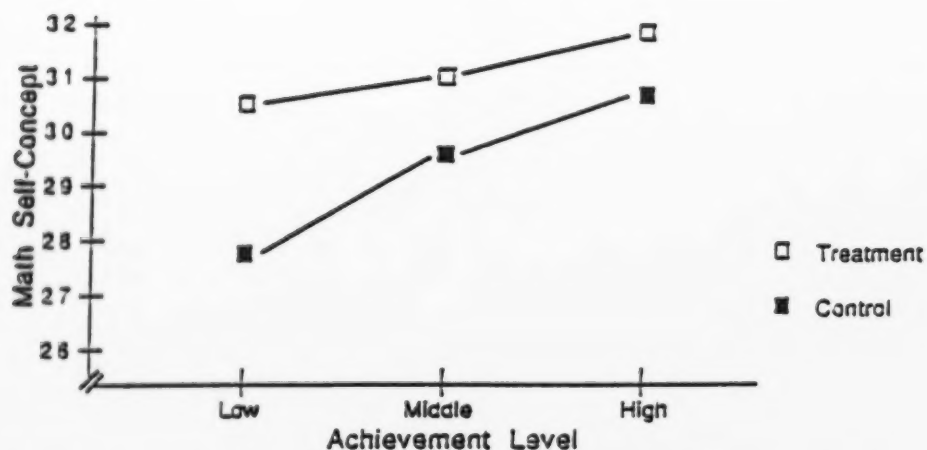


Figure. Adjusted treatment and control posttest math self-concept mean ratings by achievement level at pretest.

That is, the low-achiever's adjusted means show more of a difference between treatment/control groups than do the high students. No statistically significant interaction was found, however, which is consistent with the graphical analysis.

Considering the relatively large n-sizes (treatment--171; control--166), a small difference between groups can result in a statistically significant result. To examine the practical significance of the difference between treatment group means, a mean difference effect size between the Mastering Fractions group mean and the control group mean was computed. The standard deviation used in this calculation was the untreated groups standard deviation of the ratings at pre- and posttest. The resulting covariance adjusted mean difference effect size for treatment/control groups is +0.22, indicating a small positive effect related

to treatment group membership. The obtained value is near the lower limit of the Joint Dissemination Review Panel figure of 0.25 (Tallmadge, 1977).

However, the +0.22 value may be an overestimation. The pretest math self-concept ratings for the treatment group are higher than those of the control group. The ANCOVA procedure underadjusts the posttest ratings in this situation, which may make the statistical test liberal, and may provide an inflated estimate of the effect size. Kenny (1979) has argued that the most appropriate analysis in a quasi-experimental study is that of raw mean gain differences, assuming that standard deviations remain consistent from pre- to posttest.

The standard deviations in Table 5 above are relatively consistent, as indicated by an F-max test. When the raw gain ratings are computed from the tabled values, a standardized mean difference effect size of +0.12 ($t = +1.29$; probability = +0.2; in favor of the treatment group data) is obtained. However, the mean math self-concept ratings are relatively near the instrument ceiling. This condition may be suppressing the gains, making the raw-gain-rating effect size conservative. The true value of the effect size for this type of analysis probably falls between the two values computed, the educational effect should be considered small.

Another indication of educational significance can be derived by simple examination of the figure and Table 7,

above. The posttest adjusted mean rating for the treatment group low achievers is at essentially the same level as the mean for untreated high achievers.

Analyses using the "class" as the unit of analysis.

When the class is considered the unit of analysis, the degrees of freedom are reduced substantially ($n = 337$ to $n = 17$). The probability level associated with the mean differences between the experimental groups increases from .002 to .047. The mean difference effect size is larger than when considering the student as the unit of analysis, primarily because of the substantially reduced standard deviation. Becker & Hedges (1984) and Scruggs, White, & Bennion (1986) call attention to this fact, but do not indicate that the resulting effect size is an overestimate. However, the reader is cautioned that the results of this analysis should only be compared with other studies where the class is used as the unit of analysis.

Table 8 shows the unadjusted means for the analysis. Note that classes are not broken down by achievement level; the analyses were conducted as a one-way ANCOVA using the two-level factor--experimental condition. Also note that slight differences in the means between the "student" and "class" analyses are because of rounding in the calculations (compare Tables 5 and 8).

Table 9 shows a summary of the ANCOVA. The covariance adjusted means are 31.18 for the treatment group and 29.45 for the control group. The resulting standardized mean

Table 8

Unadjusted Raw Means, Standard Deviations, and N-sizes for the Pre- and Posttest Math Self-Concept Measures; Class as the Unit of Analysis

Experimental Group	Pretest		Posttest		n-size
	$\bar{X}$	SD	$\bar{X}$	SD	
Treatment	30.93	2.21	31.67	1.59	9
Control	28.82	1.60	28.95	1.72	8

Table 9

Analysis of Covariance for Posttest Math Self-Concept using Pretest Math Self-Concept as the Covariate--Class as the Unit of Analysis

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F	Significance of the F
Within cell	28.22	14	2.02		
Regression	12.70	1	12.70	6.30	0.025
Constant	16.78	1	16.78	8.33	0.012
RGROUP*	9.54	1	9.54	4.73	0.047

* RGROUP = Treatment group membership

difference effect size is +0.86. As in the previous analyses, the regression slopes were not statistically significantly non-parallel.

A t test of gain scores yielded a t value of 0.74 in favor of the treatment group; and a directional probability of 0.23. The mean difference effect size is 0.31 in favor of the treatment group.

Conclusions and Recommendations

Self-concept is a specific area of the affective domain which appears to have a small, reciprocal and supportive relationship with achievement. The Direct Instruction Model, designed to enhance achievement, incorporates instructional- and message-design components within a general direct-instruction framework. The Mastering Fractions program is designed to be consistent with the Direct Instruction Model, and is videodisc-based. At the same time, it was designed to be compatible with the public school environment. Mastering Fractions represents a relatively well-defined treatment. When it is in use, direct interactions between the program and the students fill about 75% of a daily math period. The teacher can concentrate on diagnostics and individualized help.

Because the Mastering Fractions program is based solidly on the Direct Instruction Model, and because it represents a well-defined, replicable treatment, it was seen as a particularly appropriate intervention to use in order to study the relationship of direct instruction with self-concept. Hypotheses for the study were derived from past research on self-concept and achievement.

Findings

Program implementation in the present study varied across classrooms. During in-service training teachers were

given techniques and recommendations from the Mastering Fractions instructor's manual. However, because of classroom peculiarities, teachers made observable modifications to the program.

Further, when classrooms were classified as having high, average, or low levels of implementation, we found that students whose teachers used techniques that were more consistent with designer recommendations had significantly higher adjusted posttest scores than those whose teachers used techniques less consistent with program recommendations. The effect persisted when achievement score variance at pretest was taken into account statistically. There remained a 20 percentage point spread between the class mean of the high-implementation classes and that of the low-implementation classes.

Given that differences in fractions achievement were evident, we examined self-concept using Marsh's instrument. The Marsh-Shavelson model of self-concept is hierarchical. Thus a program which has some effect on a specific academic subject-level self-concept, may have little or no effect on a higher level of self-concept. This appears to have been the case in the present study. Although the treatment/control differences were small, academic self-concept changed about equally for the two groups, while there were larger differences in math self-concept.

Regarding differences in mathematics self-concept, when the student is used as the unit of analysis, an analysis of

covariance (pretest math self-concept as the covariate) resulted in a statistically significant $F = 8.41$ ($p = 0.004$) for treatment- or control-group membership, but a relatively small standardized mean difference effect size of $+0.22$. But when raw gain score analyses were performed, a more conservative effect size of $+0.12$ resulted. The results of this analysis provide evidence that Mastering Fractions, an achievement-oriented program, produced a small gain in self-concept, while students in a more traditional teaching approach realized no gain.

On the other hand Peterson (1979a), in her review of open (i.e., self-concept emphasizing) educational methods, found an average effect size of $+0.16$ on self-concept when those methods were compared to traditional instruction. She found that the average effect on achievement test scores was about -0.14 . If the results of the present study are valid, then a specific form of direct instruction, used in public school classes, apparently resulted in differences in self-concept at about the same level as open methods, and major gains in content-appropriate test scores. Peterson (1979a) commented that for many of the studies she examined it was not possible to compute an effect size. Further, the level of analysis used in the primary studies was not available. In the present study (as in Corno et al., 1981) the choice of the unit of analysis made a sizeable difference in the results.

When the class was chosen as the unit of analysis, the effect size for the ANCOVA data was a +0.86 and for the raw gain score analysis -- +0.31. If Peterson's average effect size is based on the student as a unit of analysis, then the use of the Mastering Fractions program resulted in comparable self-concept changes to the average open approach. If, on the other hand, some or all of the studies Peterson reviewed used the class as the unit of analysis, there appears to be an advantage to the use of Mastering Fractions in terms of average differences in self-concept (ES-raw gain = +0.31; ES-ANCOVA = +0.86) and in terms of efficiency and effectiveness in attaining achievement differences (ES-domain specific achievement = +5.9). When the control group pre- to posttest changes in self-concept were considered, under the hypothesis-one results, it was found that the mean raw gain was -0.04. This may add an additional dimension to the interpretation of the results. As Marsh (1988) pointed out, math self-concept ratings begin to level and then to decline with age during late preadolescence; self-ratings then tend to stabilize in late adolescence. If this is the case, then the students who participated in the Mastering Fractions program may be showing a tendency to resist this decline.

It is important to note that these findings cannot be generalized to all forms of direct instruction at this time. Kelly (1986), for example, noted that there are observable differences between active teaching with a basal text and

the Mastering Fractions instructional design model, even though both are forms of direct instruction.

There were also small, but statistically significant differences in math self-concept across student-achievement levels ($F = 4.09$; $p = 0.018$). The difference between posttest math self-concept mean ratings between the treatment and control low-achieving students was slightly greater than that of the higher-achieving students.

Considering the low achievers, their posttest adjusted mean math self-concept rating (30.49) is very close to that of the untreated high achievers' adjusted posttest mean rating (30.79). That is, the Mastering Fractions program provides an environment where students who normally do not succeed in math can achieve in math and can realize a fairly high level of self-concept.

There was, however, no statistically significant interaction present. That is, high-achieving students did not lose self-concept while low-achieving students gained in self-concept. The results for this particular direct instruction program do not support suggestions by Ebmeier and Good (1979a, 1979b) and Peterson (1979b): that high-achieving students may suffer self-concept losses under direct instructional strategies.

Internal Validity

Selection may be a threat to the internal validity of this study. Even though efforts were made to recruit

similar treatment and control classes (Campbell & Stanley, 1963), there were pretest differences between the treatment groups by grade, age, math self-concept and achievement scores. Analyses of covariance and gain score analyses were both used in an effort to avoid misinterpretation of the data.

Campbell and Stanley (1963) state that the regression effect is a sure thing if there are pretest differences. In this study, however, the results are not in the same direction as would be predicted by the differences at pretest. That is, the treatment group began the study with higher math self-concept mean ratings than the control group, and yet reported even higher ratings at posttest.

As a check for a possible intrasession history effect, teachers and site staff were also asked whether any unusual events or circumstances had existed during the fall quarter which might affect the math self-concept of the control group students. They indicated that nothing unusual occurred.

Because the math SDQ scales are self-report, they are subject to fakability (Hopkins & Stanley, 1981). Instrument administration protocols developed by Marsh and associates were used to minimize fakability. In addition the Mastering Fractions program was intense for the students, and over the course of the study they began to take testing in a rather matter-of-fact way. By the end of the research project the students would probably not have been worried about whether

their responses were socially desirable. Consequently, it is likely that a social desirability effect was not present.

Generalizability of Findings

Because this study was carried out in the public schools, it was not possible to randomly select students from the national population nor to randomly assign them to classes. Given these conditions, the present study was designed to check the consistency of sample measurements against more established research findings and against national census data.

Consistency with National Norms

The sample consisted of 337 students in upper-elementary grades from three school districts. There was some variability in pretest achievement across sites and across classes. That is, whether or not students were assigned randomly to their classes by school authorities, there existed differences in achievement by class. This is probably true across the nation.

The average age and gender statistics are roughly the same as those of the 1986 census data available from the U.S. Department of Education. The only major difference in measured demographics was that the percentage of students of Hispanic origin was three times the national figure. Students of Hispanic origin did not dominate the treatment or the control group. The differences in demographics did

not appear to have a major effect on the pretest achievement and self-concept means for the sample.

If the sample in this study had not been as large as it was ($n = 337$), and had not been spread over 17 classes in 3 different districts, it is doubtful that the overall statistics would have been as similar to past census information as they were. Judging from the demographic information available, the sample is fairly representative of the national population of upper elementary school students.

Consistencies with Past Findings

Pretest achievement and self-concept measurement results were similar to those of earlier instrument standardization studies. For example, correlations of pretest norm-referenced achievement and math self-concept are consistent with values found by Marsh and associates ($r = +0.31$).

Postmeasures of the control group showed little unexplained change from the pretest values. That is, achievement scores rose as they should over a six-month period; untreated self-concept in mathematics remained essentially stable over that time. Pre- to posttest correlations and correlations of internal consistency are also similar to those in previous research. Achievement test scores are logically consistent with treated and untreated conditions. The gains in achievement as measured

by a standardized math scale indicate that the students in both the treatment and control groups steadily gained in math achievement during the study.

The consistencies found provide some evidence that the study samples are at least similar to those in previous research, and that the instruments are reliable across time and across study conditions.

Recommendations

The Mastering Fractions program appears to provide substantial gains on a criterion-referenced test of fractions. However, the use of criterion tests in this study limited the statistical interpretations of the data. Because of a pretest floor effect, there was little variability in scores at that time. Thus, pretest comparisons and correlations were constrained. In future research regarding the program, a fractions-oriented test with adequate ceiling and floor would provide better comparative data. A ceiling effect may actually have suppressed changes in the SDQ ratings.

Findings in this study provide some evidence that a specific direct instruction program (Mastering Fractions) has a positive effect on self-concept for low achievers and high achievers. Those effects are at least comparable to the average effects of open educational strategies, and may be considerably larger.

All effect sizes were positive and ranged from small to moderate in size. As there is probably no perfect statistical model for any given set of data (Kenny, 1979), future research reports in this literature should include careful considerations of the most appropriate analysis model, but include multiple analyses as recommended by Corno et al. (1981).

The Mastering Fractions program does not address areas of the curriculum outside mathematics, nor is it or programs like it, available to the student throughout his or her school career. It is possible that more pronounced differences would be found if the program were available over longer periods of time. In the past researchers have found self-concept to be relatively stable over time (Marsh, 1988), yet intense, long-duration programs (e.g., Follow Through) have apparently resulted in changes in self-concept.

Peterson found small effects on self-concept for open educational strategies. But, the term "open educational strategies" is operationally ill-defined. It may be necessary to carefully tease out the effective components of open methods, as Giaconia and Hedges (1982) are attempting to do; and then to combine them with effective achievement-enhancing components manifest in strategies found in the Direct Instruction Model (see Dawson, 1987). It could be that more student-directed work, as opposed to teacher-directed work is necessary to build solid self-concept, as

per Bandura's suggestions. At this time, however, considering the achievement efficiency of this Direct Instruction Model-based program, the use of the Mastering Fractions program to increase both achievement and self-concept is supported.

References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84, 191-215.
- Becker, B. J., & Hedges, L. V. (1984). Meta-analysis of cognitive gender differences: A comment on an analysis by Rosenthal and Rubin. Journal of Educational Psychology, 76 (1), 583-587.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago, IL: Rand McNally.
- Corno, L., Mitman, A., & Hedges, L. (1981). The influence of direct instruction on student self-appraisals: A hierarchical analysis of treatment and aptitude-treatment interaction effects. American Educational Research Journal, 18(1) 39-61.
- Dawson, M. M. (1987). Beyond ability grouping: A review of the effectiveness of ability grouping and its alternatives. School Psychology Review, 16(3), 348-369.
- Ebmeier, H., & Good, T. L. (1979a). An investigation of the interactive effects among student types, teacher types, and instructional types on the mathematics achievement of fourth grade students. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 171 549)
- Ebmeier, H., & Good, T. L. (1979b). The effects of instructing teachers about good teaching on the mathematics achievement of fourth grade students. American Educational Research Journal, 16(1), 1-16.
- Engelmann, S. (1980). Direct instruction. Englewood Cliffs, NJ: Educational Technology Publications.
- Giaconia, R. M., & Hedges, L. V. (1982). Identifying features of effective open education. Review of Educational Research, 52(4), 579-602.
- Hasselbring, T., Sherwood, B., & Bransford, J. (1986). An evaluation of the mastering fractions level-one instructional videodisc program. Nashville, TN: The Learning Technology Center, George Peabody College of Vanderbilt University.

- Hopkins, K. D., & Stanley, J. C. (1981). Educational and Psychological measurement and evaluation. (6th ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Kelly, B. F. (1986, August). The relative effectiveness of instructional design features in two fractions curricula with mildly handicapped and remedial high school students. (Doctoral dissertation, University of Oregon).
- Kenny, D. A. (1979). The non-equivalent control group design. In D. A. Kenny, Correlation and causality, (pp. 206-226). New York, NY: John Wiley & Sons.
- Lowry, W. H. (1989). The effects of a direct instruction program in fractions on academic and mathematics self-concept. (Doctoral dissertation, Utah State University).
- Marsh, H. W. (1988). Self-description questionnaire-I: SDQ manual and research monograph. San Antonio, TX: Psychological Corporation.
- Marsh, H. W., Parker, J. W., & Smith, I. D. (1983). Preadolescent self-concept: Its relation to self-concept as inferred by teachers and to academic ability. British Journal of Educational Psychology, 51(1), 60-78.
- Marsh, H. W., Relich, J. D., & Smith, I. D. (1983). Self-concept: The construct validity of interpretations based upon the SDQ. Journal of Personality and Social Psychology, 45(1), 173-187.
- Marsh, H. W. Smith, I. D. (1982). Multitrait-multimethod analyses of two self-concept instruments. Journal of Educational Psychology, 74(3), 430-440.
- Marsh, H. W., & Smith, I. D. (1987). Cross-national study of the structure and level of multidimensional self-concepts: An application of confirmatory factor analysis. Australian Journal of Psychology, 39(1), 61-77.
- Marsh, H. W., Smith, I. D., & Barnes, J. (1981). Multitrate-multimethod analyses of the Self Description Questionnaire: Student-teacher agreement on multidimensional ratings of the student self-concept. Sydney, Australia: The University of Sydney.
(ERIC Document Reproduction Service No. ED 211 581)

- Marsh, H. W., Smith, I. D., & Barnes, J. (1983). Multitrait-multimethod analysis of interpretations based upon the SDQ: Student-teacher agreement on multidimensional ratings of student self-concept. American Educational Research Journal, 20(3), 333-357.
- Marsh, H. W., Smith, I. D., & Barnes, J. (1984). Multidimensional self-concepts: Relationships with inferred self-concepts and academic achievement. Australian Journal of Psychology, 36, 367-387. (ERIC Document Reproduction Service No. ED 242 786)
- Marsh, H. W., Smith, I. D., Barnes, J., & Butler, S. (1983). Self concept: Reliability, stability, dimensionality, validity, and the measurement of change. Journal of Educational Psychology, 75(5), 772-790.
- Peterson, P. L. (1979a). Direct instruction reconsidered. In P. L. Peterson & H. J. Walberg (Eds.), Research on teaching: Concepts, findings and implications. (pp. 57-69). Berkley, CA: McCutchan Publishing.
- Peterson, P. L. (1979b, October). Direct instruction: Effective for what and for whom? Educational Leadership, pp. 46-48.
- Psychological Corporation, The (1988). The Self-Description Questionnaire-I. San Antonio, TX.
- Rosenshine, B. V. (1979). Content, time, and direct instruction. In P. L. Peterson & H. J. Walberg (Eds.), Research on teaching: Concepts, findings and implications (pp. 28-56). Berkley, CA: McCutchan Publishing.
- Scruggs, T. E., White, K. R., & Bennion, K. (1986). Teaching test-taking skills to elementary-grade students: A meta-analysis. The Elementary School Journal, 87(1), 70-82.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. Review of Educational Research, 46 (3), 407-441.
- Systems Impact, Inc. (1986). Mastering fractions [Videodiscs and Print]. Washington, DC: Author.
- Tallmadge, G. K. (1977). Ideabook: The joint dissemination review panel. Washington, DC: U. S. Office of Education.